ON THE FOUNDATIONS OF LOGIC AND ARITHMETIC.[1]

WHILE to-day in researches on the foundations of geometry we are essentially agreed as to the procedures to be adopted and the ends to be sought, it is quite otherwise with the inquiry concerning the foundations of arithmetic: here even yet the most diverse notions of the investigators stand bluntly opposed to each other.

The difficulties in the founding of arithmetic are partly indeed of a different character from those which were to be overcome in the founding of geometry.

In the examination of the foundations of geometry it was possible to leave aside certain difficulties of a purely arithmetical nature; in the founding of arithmetic, however, the appeal to another basal science seems unallowable.

I shall show the essential difficulties in the founding of arithmetic most clearly by subjecting to a brief critical discussion the views of individual investigators.

L. Kronecker, you know, saw in the concept of the whole number the true foundation of arithmetic; he formed the conception, that the integer, and that too as a general notion (parameter value), is directly and immediately given; thereby he was prevented from recognizing, that the idea of the whole number needs and is susceptible of a foundation. In so far I would designate him as a *dogmatist*: he takes the integer with its essential properties as dogma and makes no attempt to get behind it.

---

[1] Translated by George Bruce Halsted.

H. Helmholtz represents the standpoint of the *empiricist;* the standpoint of pure experience, however, seems to me to be refuted by the objection, that from experience, that is, through experiment, can never be gotten the possibility or the existence of an indefinitely great number. For the number of the things which are object of our experience, even though it is great, lies nevertheless below a finite limit.

E. B. Christoffel and all those opponents of Kronecker's, who, led by the correct feeling, that without the concept of the irrational number the whole of analysis would remain doomed to unfruitfulness, seek, by finding out "positive" properties of this concept or through like means to save the existence of the irrational number, I would designate as *opportunists.*

In my opinion, however, they have not succeeded in reaching a real refutation of the Kronecker standpoint.

Among the scientists who have penetrated more deeply into the essence of the whole number, I may mention the following:

G. Frege sets himself the problem of founding the laws of arithmetic by means of *logic*, this taken in the usual sense. He has the merit of having rightly apprehended the essential properties of the concept of the whole number as well as the significance of the inference by complete induction. Inasmuch as he, however, true to his plan, takes this also among others as axiom, that a concept (an aggregate) is defined and immediately available, provided only it be determined for every object, whether it falls under the concept or not, and also in doing this subjects the concept "every" to no restriction, he exposes himself to just those paradoxes of the theory of aggregates, which lie, for instance, in the concept of the aggregate of all aggregates and which, it seems to me, show that the conceptions and means of investigation of logic, taken in the usual sense, are not adequate to the rigorous requirements set up by the theory of aggregates.

The avoidance of such contradictions and the clearing up of those paradoxes is rather from the very outset to be fixed upon as a chief aim in researches on the number concept.

R. Dedekind has clearly perceived the mathematical difficulties

in the founding of the number concept and in most ingenious fashion first supplied a construction of the theory of whole numbers.

I would, however, designate his as a *transcendental* method in so far as he conducts his proof for the existence of the infinite in a way, whose fundamental idea is indeed used in like manner by philosophers, but which because of the unavoidable contradiction of the concept therein employed of the totality of all things, I cannot acknowledge as allowable and sure.

G. Cantor has perceived the above-mentioned contradiction and has given expression to this perception by distinguishing between "consistent" and "inconsistent" aggregates. Inasmuch as he, however, in my opinion sets up no sharp criterion for this distinction, I must designate his conception on this point as one which still leaves room for the *subjective* judgment and therefore affords no objective certainty.

I am of the opinion all the difficulties touched upon can be overcome and we can attain to a rigorous and entirely satisfactory founding of the number concept, and that by a method, which I would call *axiomatic,* whose fundamental idea I wish briefly to develop in what follows.

Arithmetic is indeed designated as a part of logic and it is customary to presuppose in founding artithmetic the traditional fundamental principles of logic.

But on attentive consideration we become aware, that in the usual exposition of the laws of logic certain fundamental concepts of arithmetic are already employed, for example the concept of the aggregate, in part also the concept of number.

We fall thus into a vicious circle and therefore to avoid paradoxes a partly simultaneous development of the laws of logic and arithmetic is requisite.

In the brief space of a lecture I can merely indicate how I conceive of this common construction. Therefore I ask indulgence if I succeed only in giving you a rough notion of what direction my researches are taking. Moreover, for the sake of being more easily understood, I shall employ the ordinary speech "in words"

and the laws of logic therein indirectly expressed, more than would be desirable in an exact construction.

Let an object of our thinking be called a *thought-thing* or briefly a *thing* and designated by a symbol.

Let us take as the basis of our consideration first of all a thought-thing 1 (one).

The taking of this thing together with itself respectively two, three or more times, as:

$$11,\ 111,\ 1111,$$

we designate as *combinations* of the thing 1 with itself; in like manner any combinations of these combinations, as:

$$(1)(11),\ (11)(11)(11),\ ((11)(11))(11),\ ((111)(1))(1)$$

are in turn called combinations of this thing 1 with itself.

The combinations likewise are designated merely as things and then in distinction to this the fundamental thought-thing 1 as *simple* thing.

We adjoin now a second simple thought-thing and denote it by the symbol $=$ (equal). We consider now in turn the combinations of these two thought-things, as:

$$1=,\ 11=,\ \ldots (1)(=1)(===),\ ((11)(1)(=))(==),\ 1=1,$$
$$(11)=(1)(1).$$

We say, the combination a of the simple things 1, $=$ *differs* from the combination b of those things, if they, as regards the mode and sequence of the combination, or the choice and participation of the things 1, $=$ themselves, deviate in any way from one another, that is if a and b are not *identical* with each other.

Now let us think the things 1, $=$ and their combinations as somehow divided into two classes, *the class of the existent* and *the class of the non-existent*: everything which belongs to the class of the existent, differs from everything which belongs to the class of the non-existent. Every combination of the two simple things 1, $=$ belongs to one of these two classes.

If a is a combination of the two fundamental things 1, $=$, then we designate also by a the *statement*, that a belongs to the class

of the existent, and by $\bar{a}$ the *statement,* that a belongs to the class of the non-existent. We designate a as a *true* statement, if a belongs to the class of the existent; on the other hand let $\bar{a}$ be called a *true* statement, if a belongs to the class of the non-existent.

The statements a and $\bar{a}$ constitute a *contradiction.*

The composite from two statements, A, B, represented in symbols by

$$\mathrm{A}|\mathrm{B},$$

in words: "from A follows B" or "if A is true, B also is true" is likewise called a statement and then A is called the *hypothesis,* B the *conclusion.*

Hypothesis and conclusion may themselves in turn consist of several statements A_1, A_2, respectively B_1, B_2, B_3 and so forth, in symbols:

$$\mathrm{A}_1 \,\&\, \mathrm{A}_2 \mid \mathrm{B}_1 \text{ o. } \mathrm{B}_2 \text{ o. } \mathrm{B}_3,$$

in words: "from A_1 and A_2 follows B_1, or B_2, or B_3" and so forth.

In consequence of the symbol o. (or) it would be possible, since negation is already introduced, to avoid the symbol | ; I use it in this lecture merely to follow as much as possible the customary word-speech.

We will understand by A_1, A_2,... respectively those statements which—to be brief—result from a statement $\mathrm{A}(x)$ when in place of the "arbitrary" x we take the thought-things 1, = and their combinations; then we write the statements

$$\mathrm{A}_1 \text{ o. } \mathrm{A}_2 \text{ o. } \mathrm{A}_3,\ldots \text{ respectively } \mathrm{A}_1 \,\&\, \mathrm{A}_2 \,\&\, \mathrm{A}_3,\ldots$$

also, as follows:

$\mathrm{A}(x^{(o)})$, in words "at least for one x"

respectively $\mathrm{A}(x^{(\&)})$, in words "for every single x,"

in this we see merely an abbreviated way of writing.

We make now from the fundamental two things 1, = the following statements:

1. $x = x$
2. $[x = y \,\&\, w\ (x)] \mid w(y)$.

Therein x (in the sense of $x^{(\&)}$) means each of the two funda-

mental thought-things and every combination of them; in 2. y (in the sense of $y^{(\&)}$) is likewise each of those things and each combination, furthermore $w(x)$ an "arbitrary" combination, which contains the "arbitrary" x, (in the sense of $x^{(\&)}$); the statement 2. reads in words:

From $x = y$ and $w(x)$ follows $w(y)$.

The statements 1., 2. form the *definition of the concept* $=$ (equal) and are in so far also called *axioms.*

If one puts in place of the arbitraries x, y in the axioms 1., 2. the simple things 1, $=$ or particular combinations of them, there result particular statements, which may be called *inferences* from those axioms.

We consider a series of certain inferences of such a sort, that the hypotheses of the last inference of the series are identical with the conclusions of the preceding inferences.

Then if we take the hypotheses of the preceding inferences as hypothesis and the conclusion of the last inference as conclusion, there results a new statement, which in turn may be designated as an inference from the axioms.

By continuation of this deduction-process we may obtain further inferences.

We select now from these inferences those which have the simple form of the statement a (affirmation without hypothesis), and comprehend the things a so resulting in the class of the existent, while the things differing from these may belong to the class of the non-existent.

We recognize, that from 1., 2. only inferences of the form $\alpha = \alpha$ ever arise, where α is a combination of the things 1, $=$.

The axioms 1., 2. in their turn as regards the partition in question of the things into the two classes are also fulfilled, that is true statements, and because of this property of the axioms 1., 2. we designate the concept $=$ (equal) defined by them as a concept *free from contradiction.*

I would call attention to the fact, that the axioms 1., 2. do not at all contain a statement of the form $\bar{a}$, that is a statement, accord-

ing to which a combination is to be found in the class of the non-existent.

We therefore could also satisfy the axioms by comprehending the combinations of the two simple things all in the class of the existent and leaving the class of the non-existent empty.

The partition above chosen into the two classes, however, shows better how to proceed in the subsequent more difficult cases.

We now carry the construction of the logical foundations of mathematical thinking further, by adjoining to the two thought-things u (infinite aggregate, infinite), f (following), f′ (accompanying operation) and laying down for them the following axioms:

$$3.\quad f(ux) = u(f'x)$$
$$4.\quad f(ux) = f(uy) \mid ux = uy$$
$$5.\quad \overline{f(ux) = u1}$$

Therein the arbitrary x (in the sense of $x^{(\&)}$) means each of the five now fundamental thought-things and every combination of them.

The thought-thing u may be called briefly ***infinite aggregate*** and the combination ux (for example u1, u(11), uf) an *element* of this infinite aggregate u.

The axiom 3. then expresses, that after each element ux follows a definite thought-thing f(ux), which is equal to an element of the aggregate u, namely to the element u(f′x), that is belongs likewise to the aggregate u.

The axiom 4. expresses the fact, that, if the same element follows two elements of the aggregate u, those elements also are equal.

According to axiom 5. there is no element in u, after which the element u1 follows; this element u1 may therefore be called the first element in u.

We have now to subject the axioms 1.—5. to the investigation corresponding to that before made of the axioms 1., 2.; therein it is to be noticed, that those axioms 1., 2. at the same time experience an extension of their validity, inasmuch as now the arbitraries x, y mean any combinations you please of the five fundamental simple things.

We ask again, whether certain inferences from the axioms 1.—5. make a contradiction or whether on the contrary the fundamental five thought-things 1, =, u, f, f′ and their combinations can be so distributed into the class of the existent and the class of the non-existent, that the axioms 1.—5. in regard to this partition into classes are fulfilled, that is, as regards that partition into classes, each inference from those axioms comes to be a true statement.

To answer this question, we take into account that axiom 5. is the only one which gives rise to statements of the form $\bar{a}$, that is that a combination a of the five fundamental thought-things is to belong to the class of the non-existent. Statements, which with 5. make a contradiction, must therefore in any case be of the form:

$$6.\quad f(ux^{(0)})=u1:$$

such an inference, however, can in no wise result from the axioms 1.—4.

In order to perceive this, we designate the equation, that is the thought-thing $a=b$ as a homogeneous equation when a and b are both combinations of two simple things, and also if a and b are both any combinations of three or both any combinations of four or more simple things; for example

$$(11)=(fu),\ (ff)=(uf'),\ (f11)=(u1=),\ (f1)(f1)=(1111),$$
$$[f(ff'u)]=(1uu1),\ [(ff)(111)]=[(1)(11)(11)],\ (fu11=)$$
$$=(uu11u)$$

are called homogeneous equations.

From the axioms 1. and 2. alone follow, as we have already seen, nothing but homogeneous equations, namely the equations of the form $\alpha=\alpha$. Just so axiom 3. gives only homogeneous equations if in it we take for x any one thought-thing.

Likewise axiom 4. is certain to exhibit in the conclusion always a homogeneous equation, if only the hypothesis is a homogeneous equation, and consequently only homogeneous equations can appear at all as inferences from the axioms 1.—4.

Now, however, the equation 6., which is the one to be proven, is certainly no homogeneous equation, since therein in place of $x^{(0)}$ one has to take a combination and thus the left side comes to be a

combination of three or more simple things, while the right side remains a combination of the two simple things u and 1.

Herewith is explained, as I think, the thought fundamental for the recognition of the correctness of my assertion; for the complete carrying through of the proof there is need of the idea of the finite ordinal number and certain theorems about the concept of equality as to number, which in fact at this point can without difficulty be set up and deduced: for the complete carrying through of the stated fundamental thought we have still to consider those points of view, to which I will briefly refer at the close of my lecture. (Compare V.)

The desired partition into classes results therefore, if one reckons as in the class of the existent all things a, where a is an inference from the axioms 1.—4., and considers as in the class of the non-existent all those things which differ from these, especially the things $f(ux) = u1$.

Because of the property of the assumed axioms so found, we recognize, that these never lead at all to a contradiction, and therefore we designate the thought-things u, f, f′ defined by them as concepts or operations *free from contradiction* or as *existing free from contradiction* (compatible).

As to the concept of the infinite u in particular, the affirmation of the *existence of the infinite* u thus appears justified through the above indicated exposition; for it gets now a definite meaning and a content continually to be applied later on.

The investigation just sketched makes the first case in which the direct proof of the freedom-from-contradiction of axioms has been successfully given, whereas the method heretofore usual—especially in geometry—for such proofs, that of suitable specialization or construction of examples, here necessarily fails.

That this direct proof here succeeds, is, as one sees, essentially owing to the circumstance, that a statement of the form $\bar{a}$, that is a statement, according to which a certain combination is to belong to the class of the non-existent, only appears as a conclusion in one place, namely in axiom 5.

When we translate the known axioms for complete induction into the speech chosen by me, we attain in like manner to the com-

patibility of the so increased axioms, that is to the proof of the contradiction-free *existence of the so-called smallest infinite** (that is, *of the ordinal type* 1, 2, 3, ...).

There is no difficulty in founding the concept of the finite ordinal number in accordance with the principles above set up; this is done on the basis of the axiom, that every aggregate which contains the first element of ordinal number and, in case any element belongs to it, also contains the one following this, surely must always contain the last element.

The proof of the compatibility of the axioms follows here very easily by the bringing in of an example, for instance of the number two. It is then the main point, to show, that an arrangement of the elements of the finite ordinal number is possible, such that each part-aggregate of it possesses a first and a last element—a fact, which we prove by defining a thought-thing $<$ by the axiom

$$(x < y \ \& \ y < z) \mid x < z$$

and then recognizing the compatibility of the axioms set up with the addition of this new axiom, when x, y, z mean arbitrary elements of the finite ordinal number.

By using the fact of the existence of the smallest infinite, the theorem then follows also, that for each finite ordinal number a still greater ordinal number can be found.

The principles which must be normative for the construction and further elaboration of the laws of mathematical thinking in the contemplated way, are briefly the following:

I. Arrived at a definite standpoint in the development of the theory, I may designate a further statement as true, as soon as is recognized, that it superadded as axiom to the statements already found true, gives no contradiction, that is leads to inferences, which in regard to a certain partition of things into the class of the existent and that of the non-existent are all true statements.

II. In the axioms the arbitraries—as equivalent for the concept "every" or "all" in the customary logic—represent only those

* Compare my lecture delivered before the International Congress of Mathematicians at Paris in 1900: *Mathematical Problems*, 2. The Compatibility of the Arithmetical Axioms.

thought-things and their combinations with one another, which at that stage are laid down as fundamental or are to be newly defined. Therefore in the deduction of inferences from the axioms, the arbitraries, which occur in the axioms, can be replaced only by such thought-things and their combinations.

Also we must duly remember, that through the superaddition and making fundamental of a new thought-thing the preceding axioms undergo an enlargement of their validity, and where necessary, are to be subjected to a change in conformity with the sense.

III. The aggregate is generally defined as a thought-thing m, and the combinations mx are called elements of the aggregate m, so that therefore—in opposition to the usual conception—the concept of the element of an aggregate appears only as later product of the idea of aggregate.

Exactly as the concept "aggregate" are also "correlation," "transformation," "reference," "function" thought-things for which, precisely as was done a moment ago with the concept "infinite," the suitable axioms are to be stated, and these then in the case of the possibility of the partition of the respective combinations into the class of the existent and that of the non-existent can be recognized as compatibly existing.

In I. the creative principle receives expression which in the freest application warrants us in ever new concept-building with the sole restriction of the avoidance of a contradiction. The paradoxes mentioned at the beginning of this lecture become because of II. and III. impossible; especially does this hold of the paradox of the aggregate of all aggregates not containing themselves as element.

In order to permit the perception of the far-going agreement in content of the concept of aggregate defined in III. with the usual aggregate-concept, I prove the following theorem:

At a definite stage in the development let

$$1, \ldots, a, \ldots, k$$

be the fundamental thought-things and $a(\xi)$ a combination of these, which contains the arbitrary ξ; further let $a(\alpha)$ be a true statement

(that is $a(\alpha)$ in the class of the existent): then there is sure to exist a thought-thing m of such a sort, that $a(mx)$ for the arbitrary x represents true statements only (that is $a(mx)$ always occurs in the class of the existent) and also inversely each thing ξ, for which $a(\xi)$ represents a true statement, is equal to a combination $mx^{(0)}$, so that the statement

$$\xi = mx^{(0)}$$

is true, that is the things ξ, for which $a(\xi)$ is a true statement, make the elements of an aggregate m in the sense of the above definition.

In proof we set up the following axiom: m is a thought-thing, for which the statements

$$7.\quad a(\xi) \mid m\xi = \xi$$
$$8.\quad \bar{a}(\xi) \mid m\xi = \alpha$$

are true, that is if ξ is a thing such that $a(\xi)$ belongs to the class of the existent, then must $m\xi = \xi$ hold good, otherwise $m\xi = \alpha$; adjoin this axiom to the axioms which are valid for the things

$$1, \ldots, \alpha, \ldots, k,$$

and then assume, that thereby a contradiction is produced, that is, that for the things

$$1, \ldots, \alpha, \ldots, k, m$$

perchance the statements

$$p(m) \text{ and } \overline{p(m)}$$

are at one and the same time inferences, where $p(m)$ is a certain combination of the things

$$1, \ldots, k, m.$$

Therewith 8. means in words the stipulation $m\xi = \alpha$, if $a(\xi)$ belongs to the class of the non-existent.

Whenever in $p(m)$ the thing m appears in the combination $m\xi$, replace in accordance with the axioms 7. and 8. and taking 2. into consideration the combination $m\xi$ by ξ, respectively α; if from $p(m)$ is formed in this way $q(m)$ (where now $q(m)$ no longer contains the thing m in a combination mx), then must $q(m)$ be an inference from the original fundamental axioms for

$$1, \ldots, \alpha, \ldots, k$$

and therewith also remain true if we for m take any one of these things, as for instance the thing 1.

Since the same consideration holds also for the statement $\overline{p(m)}$, there would therefore exist also at the original stage, when we take as a basis the things

$$1, \ldots, a, \ldots, k,$$

the contradiction

$$q(1) \text{ and } \overline{q(1)},$$

which cannot be—it being presupposed that the things

$$1, \ldots . k$$

exist free from contradiction. We must therefore reject our assumption, that a contradiction is produced; in other words, m exists free from contradiction which was to be proved.

IV. If we wish to investigate a definitely given system of axioms in accordance with the above principles, then we must partition the combinations of the fundamental things into the two classes, that of the existent and that of the non-existent, and in this process the axioms play the rôle of prescriptions which the partition must satisfy.

The chief difficulty will consist in making out the possibility of the partition of all things into the two classes, that of the existent and that of the non-existent.

The question of the possibility of this partition is essentially equivalent to the question, whether the inferences, which can be obtained from the axioms through specialization and combination in the previously exemplified sense, lead to a contradiction or not, *if besides are adjoined the familiar logical deduction-modes such as*

$$[(a|b) \ \& \ (\bar{a}|b)]|b$$
$$[(a \text{ o. } b) \ \& \ (a \text{ o. } c)]|[a \text{ o. } (b \ \& \ c)].$$

The compatibility of the axioms can then either be made out by showing how a peradventure contradiction must show itself as early as a preceding stage in the development of the theory, or by making the assumption, that there is a proof, which leads from the axioms to a definite contradiction, and then demonstrating, that

such a proof is not possible, that is to say contains in itself a contradiction. Thus indeed the proof just now sketched for the contradiction-free existence of the infinite amounts also to making out, that a proof for the equation 6. from the axioms 1.—4. is not possible.

V. Whenever in what precedes *several* thought-things, combinations, combinations of *manifold* sort or *several* arbitraries were spoken of, a limited number of such things ought always to be understood.

After the setting up of the definition of the finite number we are in position to take that mode of expression in its general meaning.

Also the meaning of the "any you please" inference and of the "differing" of one statement from all statements of a certain kind is now, on the basis of the definition of the finite number—corresponding to the idea of complete induction—through a recurrent procedure, capable of an exact description.

Thus also is to be conceived the complete carrying through of the proof just now indicated, that the statement

$$f(ux^{(0)}) = u1$$

differs from each statement which results through a finite number of steps as inference from the axioms 1.—4.: one has, that is, to consider the proof itself as a mathematical structure, namely a finite aggregate, whose elements are connected through statements expressing that the proof leads from 1.—4. to 6., and one has then to show that such a proof contains a contradiction and therefore does not in our defined sense exist free from contradiction.

In a way like that in which the existence of the smallest infinite can be proven, follows the existence of the assemblage of real numbers: in fact the axioms as I have set them up* for real numbers are expressible precisely through such formulas as the axioms hitherto laid down. As for that axiom which I have called the axiom of completeness, it expresses that the assemblage of real

* *Grundlagen der Geometrie,* second edition, Leipsic, 1903, pp. 24-26.

numbers in the sense of the reversible unique referability by elements contains every other aggregate whose elements likewise fulfill the preceding axioms; thus conceived the axiom of completeness also becomes a requirement expressible through formulas of the foregoing structure and the axioms for the assemblage of real numbers are qualitatively distinguished in no respect from the axioms necessary for the definition of the whole numbers.

In the perception of this fact lies, as I think, the real refutation of the conception of the foundations of arithmetic advocated by L. Kronecker and at the beginning of my lecture designated as dogmatic.

In like manner is shown, that contradiction-free existence belongs to the fundamental concepts of the Cantor theory of aggregates, in particular to the Cantor alefs.

D. HILBERT.

GÖTTINGEN.